田间玉米苗期高通量动态监测方法

张小青[1,2,3]，邵　松[1,2]，郭新宇[1,2]，樊江川[1,2*]

（1. 北京农业信息技术研究中心,北京 100097；2. 国家农业信息化工程技术研究中心/数字植物北京市重点实验室,北京 100097；3. 上海海洋大学 信息学院,上海 201306）

摘　要：目前对玉米出苗动态检测监测主要是依靠人工观测，耗时耗力且只能选择小的样方估算整体出苗情况。为解决人工出苗动态管理不精准的问题，实现田间精细化管理，本研究以田间作物表型高通量采集平台获取的高时序可见光图像和无人机平台获取的可见光图像两种数据源构建了不同光照条件下的玉米出苗过程图像数据集。考虑到田间存在环境背景复杂、光照不均等因素，在传统Faster R-CNN的基础上构建残差单元，使用ResNet50作为新的特征提取网络来对Faster R-CNN进行优化，首先实现对复杂田间环境下玉米出苗识别和计数；进而基于表型平台所获取的高时序图像数据，对不同品种、不同密度的玉米植株进行出苗动态连续监测，对各玉米品种的出苗持续时间和出苗整齐度进行评价分析。试验结果表明，本研究提出的方法应用于田间作物高通量表型平台出苗检测时，晴天和阴天的识别精度分别为95.67%和91.36%；应用于无人机平台出苗检测时晴天和阴天的识别精度分别91.43%和89.77%，可以满足实际应用场景下玉米出苗自动检测的需求。利用表型平台可获取时序数据的优势，进一步进行了玉米动态出苗检测分析，结果表明利用本模型得到的动态出苗结果与人工实际观测具有一致性，说明本研究提出的模型的具有鲁棒性和泛化性。

关键词：玉米苗期；Faster R-CNN；识别；计数；出苗动态监测

中图分类号：S24　　**文献标志码：**A　　**文章编号：**202103-SA003

引用格式：张小青, 邵松, 郭新宇, 樊江川. 田间玉米苗期高通量动态监测方法[J]. 智慧农业(中英文), 2021, 3(2): 88-99.

ZHANG Xiaoqing, SHAO Song, GUO Xinyu, FAN Jiangchuan. High-throughput dynamic monitoring method of field maize seedling[J]. Smart Agriculture, 2021, 3(2): 88-99. (in Chinese with English abstract)

1　引　言

作物出苗是影响田间管理决策的重要农艺因素，通过监测出苗情况可在正确的时间进行补种以减少损失[1]。传统上，通常通过用少数采样点进行目视观察（人工计数）来估测作物出苗数量[2]，既费时又费力，无法覆盖大的生产领域。随着图像获取方式的进步，计算机视觉技术在作物生长检测中的应用[3-6]得到了广泛的关注。获取高分辨率的特定地点作物出苗信息是实施精确田间管理的基础。同时，准确的作物出苗信息可用于了解土壤和环境对作物出苗的影响[7]。因

收稿日期：2021-03-11　修订日期：2021-05-17

基金项目：国家自然科学基金面上项目（31871519）；现代农业产业技术体系专项资金资助（CARS-02）；北京市农林科学院院改革与发展项目

作者简介：张小青（1995—），女，硕士研究生，研究方向为深度学习与图像处理。E-mail：15151935830@163.com。

*通讯作者：樊江川（1988—），男，硕士，工程师，研究方向为植物表型组学、农业遥感。电话：13381125869。E-mail：fanjc@nercita.org.cn。

此，为提高监测效率，近距离遥感技术被用来评估作物出苗情况。基于无人机图像数据进行的出苗检测方法，在各种作物中得到了广泛的应用[8]。目前使用无人机对作物出苗的研究主要集中在出苗数量和均匀性上[9]。

21世纪以来，成像技术的发展为高通量植物表型平台（Plant High-throughput Phenotypic Platform，HTPP）的开发提供了机会。1998年，比利时CropDesign公司成功开发世界上首套大型植物高通量表型平台，命名为TraitMill[10]。2016年，德国Lemna Tec公司开发了第一个产品级田间高通量植物表型平台Scanalyzer Field[11]，标志着植物表型获取技术正式走向大田测量。高通量表型平台的高通量、自动化、高精度等特点，极大地提高了作物图像采集的效率和准确性，同时也为玉米出苗动态监测提供了新的研究手段。

近年来，随着卷积神经网络的发展，深度学习已在自然语言处理[12]、图像识别[13]、视频跟踪[14]等领域超越了传统的机器学习算法。在农业领域，卷积神经网络较强的图像特征提取能力也使越来越多的研究人员将深度学习应用于作物表型信息获取[15]、遥感影像分类[16]等领域。深度学习模型已经用于从无人机、地面机器人系统收集的图像数据集中提取隐藏信息，以监测作物发育[17]和预测产量[18]。一些流行的深度学习模型AlexNet[19]、VGG[20]、GoogleNet[21]和ResNet[22]等被用于处理农业中的复杂数据。基于卷积神经网络的目标检测算法如YOLO[23]和SSD[24]一阶检测器系列和Faster R-CNN[25]两阶检测系列已被证实在对作物出苗检测上取得了不错的效果。

根据以上分析，本研究目的是利用田间作物表型平台获取的高时序、连续可见光图像以及无人机平台获取的可见光图像构建数据集，提出基于卷积神经网络（Convolutional Neural Network，CNN）的图像特征自动提取方法，实现对复杂的田间环境下玉米幼苗的快速准确识别，为玉米田间智能耕作管理提供技术储备。基于优化的Faster R-CNN模型来实现玉米出苗检测，结合非极大值抑制方法（Non-Maximum Suppression，NMS）对单位面积的玉米幼苗进行计数。依据表型平台的时序图像实现不同玉米品种的出苗动态监测，以此来判断出不同品种玉米的出苗特征及差异，并对两种平台的出苗检测精度进行验证比较。

2 材料与方法

2.1 图像采集

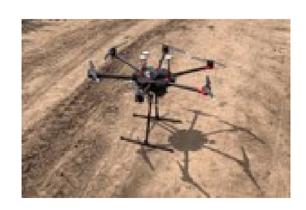
研究分别使用无人机平台和田间作物表型高通量采集平台（轨道式）搭载可见光传感器进行高通量原始图像数据获取。因此，图像采集试验分为两部分，其中试验一使用无人机平台进行图像获取，试验二使用田间作物表型高通量采集平台进行图像获取。

2.1.1 无人机可见光图像采集

试验一于2018年在北京市通州区实验基地开展（39°56′ N，116°16′ E）。使用大疆M600 Pro六旋翼无人机作为图像采集平台，搭载型号为SONY-5100的相机（采用20 mm的定制机械镜头）对试验田内玉米苗期冠层图像进行采集。考虑到小型无人机的续航能力和图像的分辨率等问题，设置无人机的飞行高度为30 m，航向重叠度为70%，旁向重叠度为65%。试验的采集时间为2018年5月8日（出苗）至2018年5月23日。考虑所获取图像的质量问题，选择在无风或者微风的时候用无人机采集数据。所获取的可见光原始图像分辨率为5280×3956 px。图1展示了无人机实物图、田间场景图以及无人机获取可见光图像。

2.1.2 表型平台可见光图像采集

试验二在北京市农林科学院试验田内开展（39°56′ N，116°16′ E）。试验二地块长15 m，宽6 m。在其上方搭建了数字植物北京市重点实验室自主研发的田间作物高通量表型平台（轨道式）。在平台的覆盖范围内种植了AD268和XD20两个玉米品种，每个品种分别以25,000株/ha，


（a）无人机数据采集装置


（b）田间场景


（c）无人机可见光原始图像

图1　无人机图像采集

Fig. 1 Image collection using UAV

40,000株/ha和55,000株/ha三种密度进行种植，共计6个小区，每个小区种植4行，共24行，行长2.5 m。两个品种玉米于2020年7月23日播种，种植后以正常水肥管理。图2为地块中种植情况示意图。

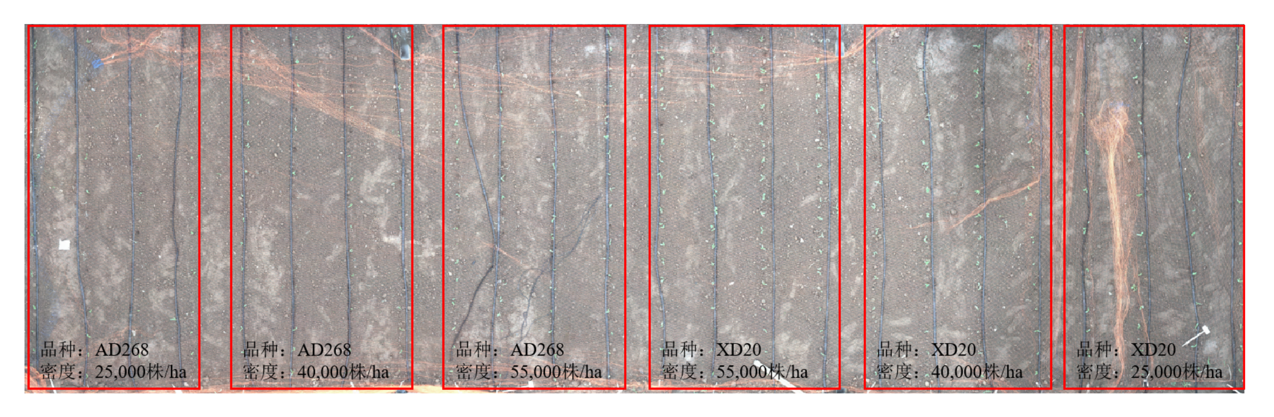


图2　田间作物高通量表型平台（轨道式）覆盖范围内玉米种植情况示意图

Fig. 2 Schematic diagram of maize planting in the coverage area of HTPP

在本次试验中，使用Microsoft，Kinect V2.0作为成像单元图像采集设备，有效像素为500万px。镜头垂直向下，采用自动曝光模式获取玉米冠层图像。图像采集时间从2020年7月26日（出苗）开始到2020年9月28日止，每天分别选择上午10点和下午3点两个时间点采集玉米出苗冠层图像，所获取的图像原始分辨率为2048×2048 px。田间作物高通量表型平台图像采集装置如图3所示。


（a）表型平台实景图


（b）图像采集设备

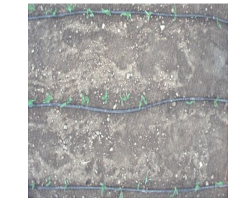

（c）表型平台可见光原始图像

图3　田间作物高通量表型平台图像采集装置

Fig.3 Field crop high-throughput phenotyping platform image acquisition device

2.2 数据集构建

2.2.1　图像预处理

综合考虑硬件条件和试验参数的需要，同时为了便于模型训练，将试验二中获取的原始图像（分辨率为2048×2048 px）缩放至960×960 px输入模型。而为了与田间作物高通量表型平台获

取的图像保持统一，对试验一中获取的无人机可见光原始图像（分辨率为5280×3956 px）同样缩放至960×960 px，采用滑动步长剪裁的方式裁剪图像。并且为了充分覆盖到遥感影像中目标的信息，裁剪时保证一定的重叠度，从而能够提高模型后期训练的效率和精度。

使用图像注释工具LabelImg对裁剪及缩放后的图像进行幼苗的边界框标注，通过最小外接矩形将样本中的玉米苗标出，记录该矩形的左上和右下的顶点坐标，并设置类别标签为seedling。图4展示了数据从缩放、裁剪到标注的整体流程。

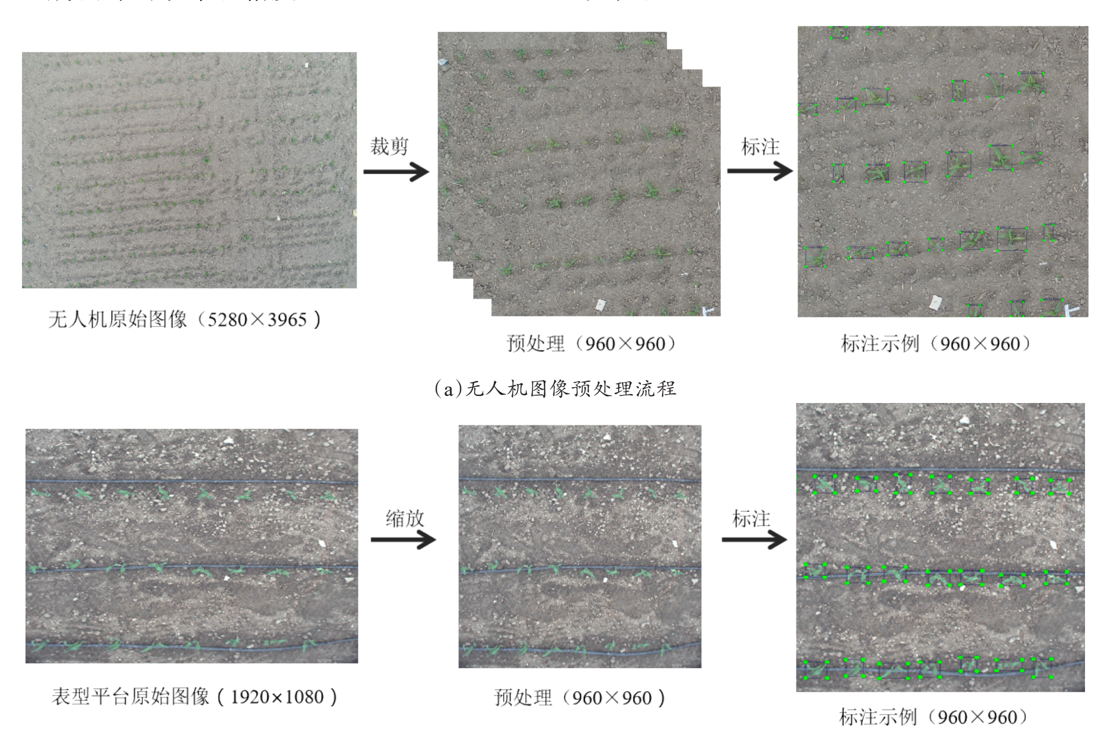


图4 无人机和田间作物高通量表型平台数据预处理流程图

Fig. 4 Data preprocessing flowchart of UAV and HTPP platforms

2.2.2 最终数据集组成

为实现玉米出苗检测的目标，本研究选取试验一中玉米刚出苗时期的无人机获取的图像作为数据集，即5月8日至5月10日，共计4天。分别选取每天原始图像数据28张，共计112张。经过滑动步长裁剪后，每张原始图像被裁剪成12张960×960 px的图像，共计1344张。考虑到图像的质量问题，手动剔除了过度曝光、质量较低的图像。经过筛选后，剩余1188张图像构成试验一数据集。将此数据集按照7:3的比例划分为训练集和测试集。对于试验二，选取玉米刚出苗时期的表型平台获取的图像作为数据集，即7月28日至8月3日，共计6天。分别选取每天原始图像数据150张，共计1200张。经过缩放后放入试验二数据集。同时由于所选取日期的天气情况不同（即同时包括了晴天和阴天），能够保证不同光照条件下模型的泛化性能。

本研究最终数据集组成如表1所示。图5展示了两种平台数据分别在不同光照条件下预处理后的图像及其标注，其中图5（a）为不同光照条件下无人机平台图像及其标注，图5（b）为不同光照条件下表型平台图像及其标注。

2.3 玉米幼苗检测网络

本研究基于卷积神经网络在图像特征提取上的优异表现，构建了基于Faster R-CNN玉米出苗

表1　数据集组成

Table 1 Composition of data set

数据来源	训练集图像数量/张	测试集图像数量/张	总数/张
无人机平台	832	356	1188
田间表型平台	840	360	1200
总数	1672	716	2388

动态检测模型，该模型以无人机和田间作物高通量表型平台采集的玉米冠层图像为数据源，对玉米田间出苗状况进行监测。玉米出苗网络的检测过程如图6所示。

2.3.1　Faster R-CNN模型框架

本研究采用Faster R-CNN深度学习模型作为基础框架，其网络结构如图7所示。

由图7可见，Faster R-CNN网络主要有四个主要内容。首先是卷积层对输入的图片，Faster R-CNN使用一组基础的conv+relu+pooling层提取图片的特征图（Feature map），用于后续候选框检测网络（Region Proposal Network）和全连接层。在RPN层中使用尺寸为3×3 px的滑动窗口遍历该特征图，以每个锚点（滑动窗口中心）为中心，设计不同的长宽比生成多尺度候选区域。ROI Pooling层在此基础上提取候选特征图（Proposal Feature Maps），得到固定尺寸的候选特征。最后经过分类和回归，输出候选区所属的类和候选区域在图中的精确位置。

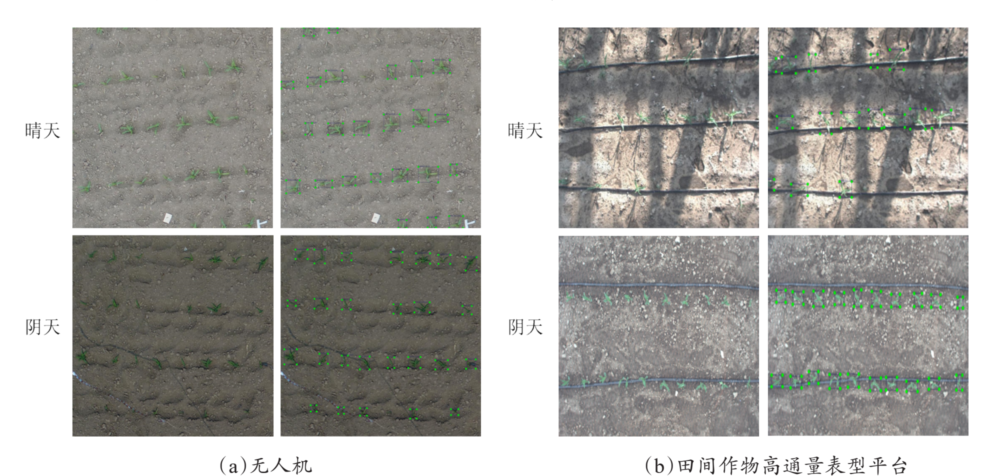
2.3.2　ResNet50特征提取网络

本研究以田间玉米冠层图像为数据，进行出苗动态检测。田间场景比较复杂，不同光照和不同采集装置得到的图像具有很大的差异，传统的Faster R-CNN网络中特征提取网络为VGG。VGG结构层次相对较浅，难以达到预期效果，因此本研究在Faster R-CNN网络基础上设计应用残差（Residual）学习单元，使用ResNet50作为特征提取网络。

研究中用到的特征提取网络ResNet 50结构如图8所示。ResNet 50共有5个编码层，第一个编码层是由卷积层和池化层构成，后面4个编码层就是4个残差块。残差结构中包括两个分支，其中，第一个分支是使用三个卷积层、批归一化层以及修正线性单元（Rectified Linear Unit，ReLu）[26]层进行特征处理，而另一个分支则是通过直接映射处理。在使用特征映射融合将这两个分支集成起来后，整个网络只需要学习输入和输出之间的差异，简化了学习的目标和难度。表2展示了特征提取网络的具体参数配置。

2.3.3　激活函数与损失函数

本研究中网络训练时选取ReLu作为激活函数，当输入值为负时，ReLu函数的输出值恒为0；当输入值为正时，ReLu函数的值与输入值相等，即：



（a）无人机　　（b）田间作物高通量表型平台

图5　不同光照条件下的可见光图像及其标注

Fig. 5 RGB images and their labels under different light conditions

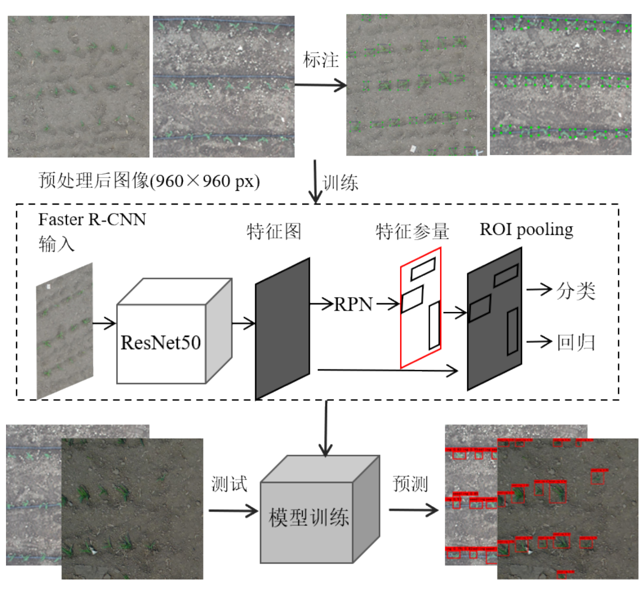


图 6 利用 Faster R-CNN 网络检测出苗流程

Fig. 6 Seedling detection process using Faster R-CNN

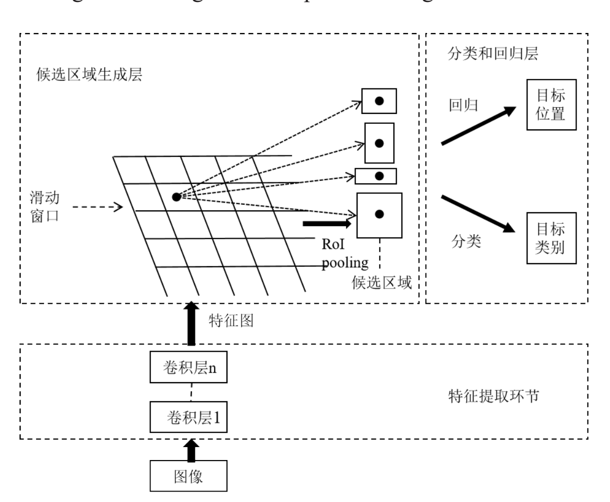


图 7 Faster R-CNN 网络框架

Fig. 7 Faster R-CNN Network framework

$$\mathrm{ReLu}(x)=\begin{cases}x & x>0\\ 0 & x\leqslant 0\end{cases} \tag{1}$$

本研究模型的损失函数主要由两部分组成，其中第一部分是分类损失[27]，其具体公式如下：

$$L=\frac{1}{N_{cls}}\sum_i (p_i, p_i^*) \tag{2}$$

其中，p_i为第i个检测框的预测值；当预测值与标注区域的重叠区域大于0.7时p_i^*=1，当预测值与标注区域的重叠区域小于0.3时p_i^*=0；N_{cls}表示网络训练过程中生成的256个检测框。

第二部分是回归损失，用于确定网络训练过程中的偏移量，其具体公式为：

$$L=\lambda\frac{1}{reg}\sum_i p_i^* L_{reg}(t_i, t_i^*) \tag{3}$$

其中，t_i表示训练阶段预测值的偏移量；t_i^*表示训练阶段相对于实际的偏移量；λ是平衡回归损失和分类损失的参数。

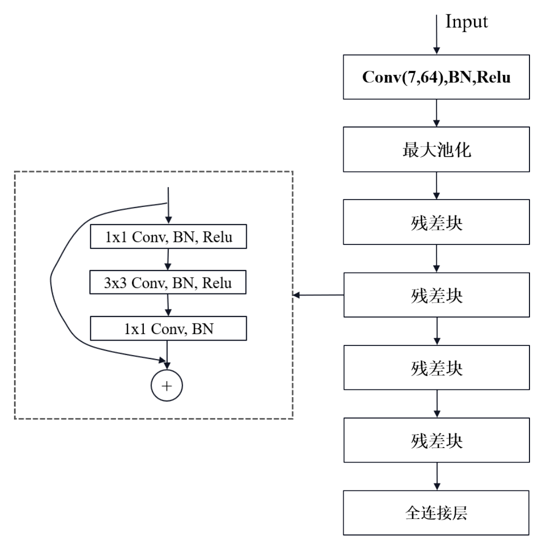


图 8 ResNet 50 网络结构图

Fig. 8 Network structure of ResNet 50

表 2 ResNet 50 参数配置

Table 2 Parameters of ResNet 50

网络层	卷积核	通道数	残差结构数量
Layer-1	卷积操作(卷积核7×7,64),Stride=2 最大池化操作(池化矩形3×3),Stride=2		
Layer-2	1×1 3×3 1×1	64 64 256	3
Layer-3	1×1 3×3 1×1	128 128 512	4
Layer-4	1×1 3×3 1×1	256 256 1024	6
Layer-5	1×1 3×3 1×1	512 512 2048	3
Layer-6	平均池化操作1×1,FC-1000,Softmax		

3 结果与分析

3.1 试验参数设置及训练过程分析

模型训练硬件平台采用 Intel（R） Xeon

（R） CPU E5-2683 v3处理器，128 GB内存，两块GTX 1070Ti显卡，以及8GB显存。软件环境在Windows 10操作系统下进行配置，语言开发环境为Python3.6、Anaconda3和TensorFlow 1.13.2。

选择随机梯度下降法（Stochastic Gradient Descent，SGD）来更新学习率，从而对Faster R-CNN网络进行训练。训练时将初始学习率设置为$1e^{-4}$，衰减率设置为0.9。为了加快训练速度和训练初期防止权值被破坏，在迭代前20个epoch时候，冻结一部分网络权值的更新，迭代20个epoch之后把所有神经网络进行解冻训练，同时将学习率调整为$1e^{-5}$，衰减率依然设置为0.9。设置网络训练的迭代次数为500个epoch，每个epoch训练1000步。训练Loss曲线如图9所示。

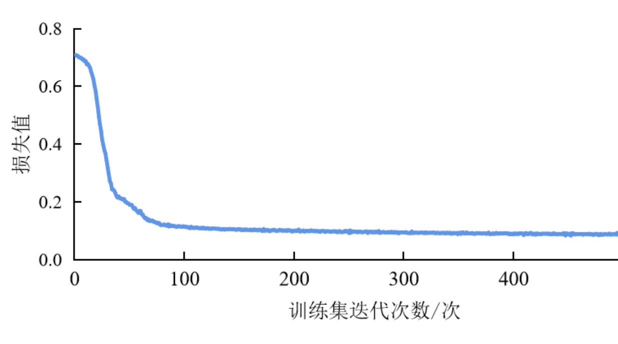


图9 Faster R-CNN模型训练Loss曲线

Fig. 9 Loss value curve of Faster R-CNN

由图9可以看出，模型在前40个批次损失快速下降，40个批次后模型的损失曲线趋于平稳。不断对模型进行训练，直到损失趋于稳定并不再变化，最终迭代到500次时，停止训练。

3.2 模型训练结果

该模型对于无人机可见光以及田间作物高通量表型平台可见光两种图像在不同光照条件下的玉米幼苗检测效果如图10所示。直观上可以看出，以ResNet50作为特征提取网络进行的幼苗检测，具有更好的检测结果，而且晴天检测结果优于阴天检测结果。表型平台图像中基本所有出苗都被有效地检测出来，而无人机图像中有很少一部分出苗的玉米没有被检测出来。

3.3 评价指标

对训练结束之后的模型进行性能判断，对于分类模型通常采用精确率（Precision）、召回率（Recall）、交并比（Intersection Over Union，IOU）作为评价指标，而平均绝对误差（Mean Absolute Error，MAE）通常被用来评价计数模型的性能。本研究采用以上4个评价指标来对模型的检测计数性能进行客观的评估。各评价指标的公式如下：

$$P = \frac{TP}{TP + FP} \tag{4}$$

$$R = \frac{TP}{TP + FN} \tag{5}$$

其中，P为Precision，表示在所有模型预测为玉米幼苗的结果中，预测正确数量占总数的比例；R为Recall，表示模型预测为玉米幼苗的数量占实际幼苗数量的比例；TP（True Positive）表示正阳性是指样本预测值为正实际也为正；FP（False Positive）表示假阳性是指样本预测值为正实际为负；FN（False Negative）表示假阴性是指样本预测值为负实际为正的。

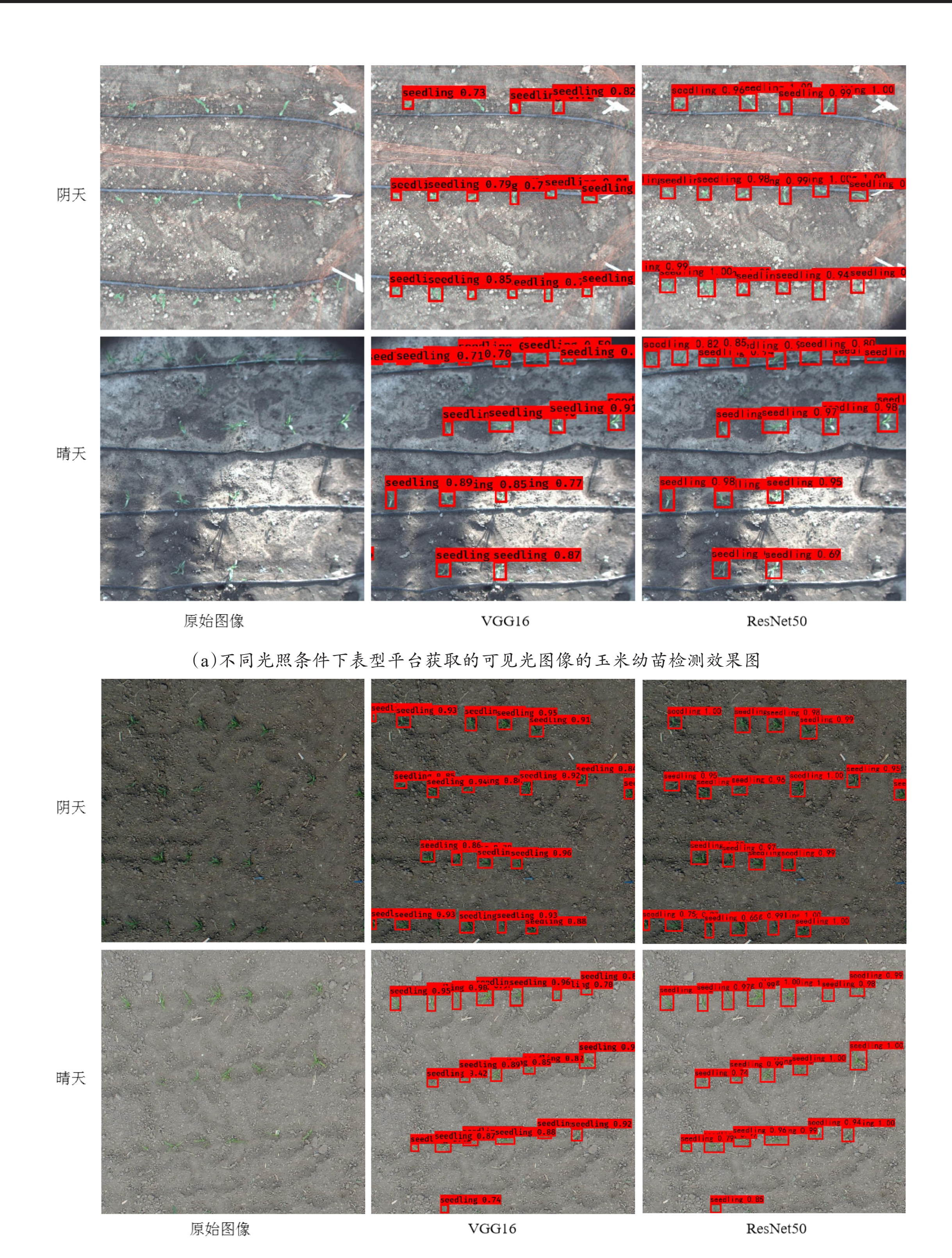
$$IOU = \frac{|A \cap B|}{|A \cup B|} \tag{6}$$

其中，IOU表示模型预测的边界框与真实标注的边界框的交集和并集的比值；A表示模型预测的边界框；B表示人工标注的边界框即为真实值。

$$MAE = \frac{1}{m} \sum_{i=1}^{m} \left| y_i - \hat{y}_i \right| \tag{7}$$

其中，$\hat{y}_i$表示预测值；y_i表示真实值；$\hat{y}_i$和y_i可以是任意形状，每个形状包含m个元素。MAE即为绝对误差的平均值，可以更好地反映预测值误差的实际情况，平均绝对误差值越小即表示模型训练的效果越好。

根据上述所列出的评价指标对无人机图像和表型平台图像进行分析，结果如表3～表5所示。图11（97页）为模型在不同平台和不同天气下输出的P-R曲线。PR曲线中将recall设置为横坐



(a)不同光照条件下表型平台获取的可见光图像的玉米幼苗检测效果图

(b)不同光照条件下无人机平台获取的可见光图像的玉米幼苗检测效果图

图 10 不同光照条件下表型平台和无人机平台获取的可见光图像的玉米幼苗检测效果图

Fig. 10 The effect of maize seedling detection on visible light images acquired by HTPP and UAV under different lighting conditions

标，precision设置为纵坐标。曲线与坐标周包围的面积为AP，AP越大，说明检测结果越好

表3　基于VGG和ResNet的Faster R-CNN在测试集输出结果对比

Table 3 Comparison of output results of Faster R-CNN based on VGG and ResNet in the test set

特征提取网络	*P*/%	*R*/%	*IOU*	*MAE*
VGG	89.42	86.81	75.04	6.2816
ResNet	91.87	90.13	82.57	5.6361

表4　本研究模型在表型平台晴天和阴天图像输出结果比较

Table 4 Comparison of the output results of the model in the images of HTPP on sunny and cloudy days

光照情况	*P*/%	*R*/%	*IOU*	*MAE*
晴天	95.67	93.36	85.57	4.3682
阴天	91.36	89.23	83.04	5.2314

表5　本研究模型在无人机晴天和阴天图像输出结果比较

Table 5 Comparison of the output results of the model in the UAV images on sunny and cloudy days

光照情况	*P*/%	*R*/%	*IOU*	*MAE*
晴天	91.43	89.46	86.01	5.3521
阴天	89.77	87.81	83.04	5.9564

由表3可以看到本研究基于ResNet特征提取网络的Faster R-CNN在整体测试集上的表现要优于传统的Faster R-CNN模型。由表4和表5可以看出本研究模型晴天时在高通量表型平台上精度可以达到95.67%，表现要优于无人机平台精度（91.43%），阴天时在表型平台和无人机平台上的识别精度也可以达到91.36%和89.77%。改进后的Faster R-CNN模型可以满足对高通量表型平台和无人机平台对出苗检测的要求。

3.4 玉米出苗动态监测

考虑到田间作物高通量表型平台能够获取每天的时序性可见光图像（图12）。本研究采用其所获取图像实现不同品种玉米的动态出苗检测。平台覆盖地块中分别以三种不同密度种植了AD268和XD20两种玉米植株，每种密度四次重复。除去保护行，共24行，行长为2.5 m即密度为25,000株/ha每行种植6棵植株，密度为40,000株/ha每行种植9棵植株，密度为55,000株/ha每行种植12棵植株。各品种玉米于7月23日同时播种。使用本研究训练的模型从7月26日（未出苗）开始对各品种玉米连续10 d的冠层图像进行出苗检测计数，直至所有植株完全出苗。统计并绘制了同一个品种不同密度的植株出苗数量的变化曲线，如图12所示。

从图12（a）可以看出AD268品种玉米植株于7月27日开始出苗，于8月4日全部出苗结束。而图12（b）的XD20品种玉米植株同样于7月27日开始出苗，三种不同密度植株在8月1日均全部出苗。为了更加直观地展示不同密度玉米植株的出苗情况，分别统计了不同种类和不同种植密度的玉米植株从第一天出苗至出苗结束所需天数，如表6所示。

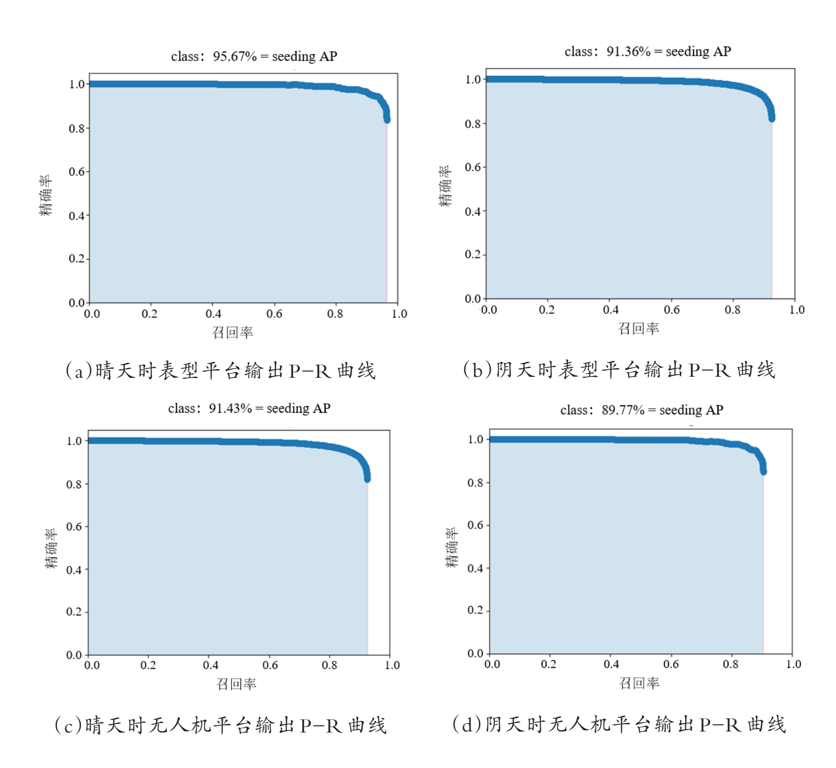
结合图12可知，XD20品种的玉米植株比AD268品种的玉米植株从开始出苗到结束所需的时间更短即出苗更整齐。而在本次试验中，在相同种植条件及灌溉条件下，密度为25,000株/ha的玉米植株从开始出苗至结束所需时间相对另两种密度玉米植株更短。利用本研究所提模型进行动态出苗检测结果与人工实测结果保持一致，说明本研究所提方法具有一定的鲁棒性和泛化性。

4　结　论

本研究以无人机平台可见光图像以及田间作物表型高通量采集平台可见光图像两种数据源构建了不同光照条件下玉米冠层图像数据集。使用Resnet 50代替VGG 16来优化Faster R-CNN模型，对所获取共计2388张图像进行了训练和测试。优化后的模型在测试集上的平均识别精度为91.87%，并对在不同平台和不同光照下的模型检测精度进行了对比。其中，高通量表型平台对出苗进行预测时，晴天和阴天的识别精度可以分别达到95.67%和91.63%。无人机平台对出苗进行预测时，晴天和阴天的识别精度可以分别达到

91.36%和89.77%。研究结果表明，应用本模型可实现在田间作物高通量表型平台上玉米幼苗检测，达到可观的效果。而在无人机平台上的检测精度却相对偏低。这是由于两种平台所获取的图像空间分辨率大小不一样、采用的预处理方式也不一样。导致处理后两种平台的图像数据所包含的玉米幼苗数量也不一致，因此在进行幼苗检测时会出现不同平台的检测精度差。

在此基础上，根据田间作物高通量表型平台所获取的时序图像对玉米出苗数量变化进行统计，分析了本试验中AD268和XD20两个品种玉米在三种不同种植密度条件下，从出苗开始至结束所需时间。结果表明，不同品种的出苗持续时间存在显著差异，XD20相比AD268出苗持续时间更短，出苗更整齐。本研究方法具备一定的准确性和科学性，同时也为使用无人机和表型平台获取玉米冠层图像进行识别计数任务及出苗动态检测提供了一定的理论依据和借鉴价值。

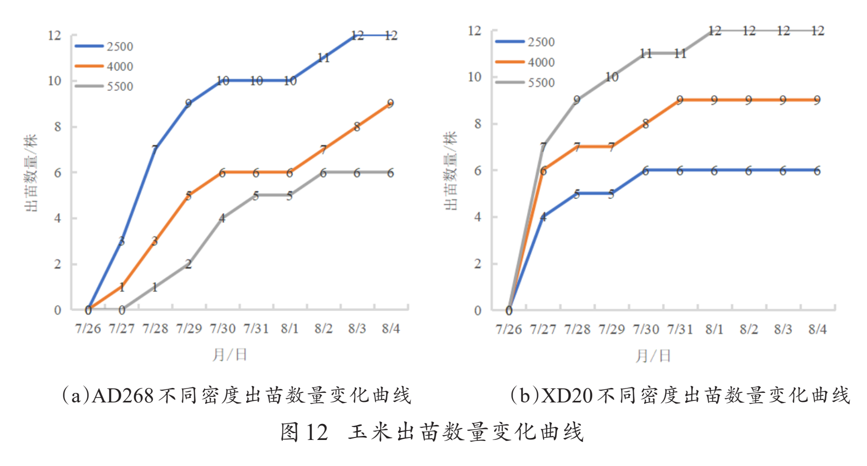


(a)晴天时表型平台输出P-R曲线　(b)阴天时表型平台输出P-R曲线

(c)晴天时无人机平台输出P-R曲线　(d)阴天时无人机平台输出P-R曲线

图11　两种表型平台在阴天和晴天时输出P-R曲线

Fig. 11 P-R curves of two phenotypic platforms on cloudy and sunny days



(a)AD268不同密度出苗数量变化曲线　(b)XD20不同密度出苗数量变化曲线

图12　玉米出苗数量变化曲线

Fig. 12 The curves of the number of maize emergence

表6　不同品种不同密度玉米植株出苗至结束所需天数统计表

Table 6 Statistics of the number of days required from emergence to end of different varieties and different densities meter

品种	种植密度/(株·ha^{-1})	出苗天数/天
AD	25,000	5
	40,000	8
	55,000	7
XD	25,000	3
	40,000	4
	55,000	5

本研究根据田间作物表型平台从玉米未出苗开始连续10天所获取的时序可见光图像，进一步应用该优化模型对不同玉米品种在不同种植密度下的出苗动态进行了监测。以上结果说明本研究建立的基于深度学习的玉米出苗检测模型不仅可以实现无人机可见光图像中玉米幼苗的准确识别与检测，也能够通过田间作物表型高通量平台实现对玉米出苗的动态连续监测。通过对两种平

台的对比，田间表型高通量平台相对于无人机表型平台所获取的图像更加精细，时序性更强，所以出苗检测的准确度更高，能够更好地为玉米播种质量和种子质量评价提供技术支撑，为田间大面积测定出苗率及产量估算提供参考。

参考文献：

[1] JIM S. New systems enhance seed and product-placement at planting[J]. Cotton Grower, 2021, 57(1): 16-18.

[2] WILES L J, SCHWEIZER E E. The cost of counting and identifying weed seeds and seedlings[J]. Weed Science, 1999, 47(6): 667-673.

[3] ZHAO W, YAMADA W, Li T, et al. Augmenting crop detection for precision agriculture with deep visual transfer learning: A case study of bale detection[J]. Remote Sensing, 2021, 13: ID 23.

[4] SAKO Y, MCDONALD M, FUJIMURA K, et al. A system for automated seed vigour assessment[J]. Seed Science and Technology, 2001, 29(3): 625-636.

[5] HOFMASTER A, FUJIMURA K, MCDONALD M, et al. An automated system for vigor testing three-day-old soybean seedlings[J]. Seed Science and Technology, 2003, 31(3): 701-713.

[6] MARCOS-FILHO J, BENNETT M, MCDONALD M, et al. Assessment of melon seed vigour by an automated computer imaging system compared to traditional procedures[J]. Seed Science and Technology, 2006, 34 (2): 485-497.

[7] FORCELLA F, ARNOLD R L B, SANCHEZ R, et al. Modeling seedling emergence[J]. Field Crops Research, 2000, 67(2): 123-139.

[8] FENG A, ZHOU J, VORIES E, et al. Evaluation of cotton emergence using UAV-based narrow-band spectral imagery with customized image alignment and stitching algorithms[J]. Remote Sensing, 2020, 12(11): ID 1764.

[9] GNADINGER F, SCHMIDHALTER U. Digital counts of maize plants by unmanned aerial vehicles (UAVs)[J]. Remote sensing, 2017, 9(6): ID 544.

[10] REUZEAU C, FRANKARD V, HATZEELD Y, et al. Traitmill™: A functional genomics platform for the phenotypic analysis of cereals[J]. Plant Genetic Resources: Characterisation and Utilisation, 2006, 4: 20-24.

[11] VIRLET N, SABERMANESH K, SADEGHI - TEHRAN P, et al. Field scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring[J]. Functional Plant Biology, 2017, 44(1): 143-153.

[12] 李寒冰 . 基于深度学习的舆情监测系统[D]. 北京: 北京交通大学, 2019.
LI H. Public opinion monitoring system based on deep learning[D]. Beijing: Beijing Jiaotong University, 2019.

[13] 李卫 . 深度学习在图像识别中的研究及应用[D]. 武汉: 武汉理工大学, 2014.
LI W. The research and application of deep learning in image recognition[D]. Wuhan: Wuhan University of Technology, 2014.

[14] 史璐璐 . 深度学习及其在视频目标跟踪中的应用研究[D]. 南京: 南京邮电大学, 2019.
SHI L. Research on deep learning and its applications in video object tracking[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2019.

[15] GRINBLAT G L, UZAL L, GLARESE M. Deep learning for plant identification using vein morphological patterns[J]. Computers and Electronics in Agriculture, 2016, 127: 418-424.

[16] 杨建宇, 周振旭, 杜贞容, 等 . 基于SegNet语义模型的高分辨率遥感影像农村建设用地提取[J]. 农业工程学报, 2019, 35(5): 251-258.
YANG J, ZHOU Z, DU Z, et al. Rural construction land extraction from high spatial resolution remote sensing image based on SegNet semantic segmentation model[J]. Transactions of the CSAE, 2019, 35(5): 251-258.

[17] HUMPHREY E J, BELLO J P, LECUN Y. Moving beyond feature design: Deep architectures and automatic feature learning in music informatics[C]// The 13th International Society for Music Information Retrieval Conference. Piscataway, New York, USA: IEEE, 2012: 403-408.

[18] NANNI L, GHIDONI S, BRAHNAM S. Handcrafted vs. non-handcrafted features for computer vision classification[J]. Pattern Recognition, 2017, 71: 158-172.

[19] KRIZHEVSKY A. One weird trick for parallelizing convolutional neural networks[EB/OL]. 2014. http://arxiv.org/abs/1404.5997v2.

[20] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. 2014. arXiv:1409.1556 [cs.CV].

[21] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with

convolutions[C]// The IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, New York, USA: IEEE, 2015: 1-9.

[22] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// The IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: IEEE Computer Society, 2016: 770-778.

[23] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// The IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, New York, USA: IEEE, 2016: 779-788.

[24] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]// European Conference on Computer Vision. Springer, Cham, Switzerland: 2016: 21-37.

[25] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.

[26] MADEC S, JIN X, LU H, et al. Ear density estimation from high resolution RGB imagery using deep learning technique[J]. Agricultural and Forest Meteorology, 2019, 264: 225-234.

[27] ZOU H, LU H, LI Y, et al. Maize tassels detection: A benchmark of the state of the art[J]. Plant Methods, 2020, 16(1): 1-15.

High-Throughput Dynamic Monitoring Method of Field Maize Seedling

ZHANG Xiaoqing[1,2,3], SHAO Song[1,2], GUO Xinyu[1,2], FAN Jiangchuan[1,2*]

(1. *Beijing Research Center for Information Technology in Agriculture, Beijing* 100097*, China*; 2. *Beijing Key Lab of Digital Plant, National Engineering Research Center for Information Technology in Agriculture, Beijing* 100097*, China*; 3. *College of Information Technology, Shanghai Ocean University, Shanghai* 201306*, China*)

Abstract: At present, the dynamic detection and monitoring of maize seedling mainly rely on manual observation, which is time-consuming and laborious, and only small quadrats can be selected to estimate the overall emergence situation. In this research, two kinds of data sources, the high-time-series RGB images obtained by the plant high-throughput phenotypic platform (HTPP) and the RGB images obtained by the unmanned aerial vehicle (UAV) platform, were used to construct the image data set of maize seedling process under different light conditions. Considering the complex background and uneven illumination in the field environment, a residual unit based on the Faster R-CNN was built and ResNet50 was used as a new feature extraction network to optimize Faster R-CNN to realize the detection and counting of maize seedlings in complex field environment. Then, based on the high time series image data obtained by the HTPP, the dynamic continuous monitoring of maize seedlings of different varieties and densities was carried out, and the seedling duration and uniformity of each maize variety were evaluated and analyzed. The experimental results showed that the recognition accuracy of the proposed method was 95.67% in sunny days and 91.36% in cloudy days when it was applied to the phenotypic platform in the field. When applied to the UAV platform to monitor the emergence of maize, the recognition accuracy of sunny and cloudy days was 91.43% and 89.77% respectively. The detection accuracy of the phenotyping platform image was higher, which could meet the needs of automatic detection of maize emergence in actual application scenarios. In order to further verify the robustness and generalization of the model, HTPP was used to obtain time series data, and the dynamic emergence of maize was analyzed. The results showed that the dynamic emergence results obtained by HTPP were consistent with the manual observation results, which shows that the model proposed in this research is robust and generalizable.

Key words: field maize; Faster R-CNN; recognition; counting; dynamic seedling detection